**What is claimed is:**

1. A computer-assisted method for determining a plurality of clusters, comprising the activities of:

   for each of a plurality of observations, obtaining a data set containing no more than one proxy value for each of a plurality of variables, each variable having a plurality of possible values;

   assigning each of the plurality of observations to one of a plurality of clusters; and

   for the plurality of observations and the plurality of variables, via cluster reassignment, maximizing a processor-determined fitness score representing a number of variables from the plurality of variables for which each observation's proxy value corresponds to a mode for that observation's assigned cluster.

2. The method of claim 1, wherein the proxy value represents a single provided value.

3. The method of claim 1, further comprising transforming a plurality of provided values for a particular variable into a single proxy value.

4. The method of claim 1, further comprising transforming a single provided value for a particular variable into a single proxy value.

5. The method of claim 1, further comprising transforming a single provided continuous value to a particular continuous variable into a single categorical proxy value.

6. The method of claim 1, wherein a portion of the data set can be bias-free.
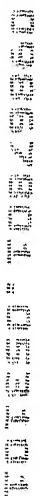
7. The method of claim 1, wherein a portion of the data set can be obtained from a panel that objectively measures behaviors.

8. The method of claim 1, further comprising calculating the fitness score.

9. The method of claim 1, further comprising calculating the fitness score by activities comprising:
   a) calculating modes of proxy values for each variable for all observations associated with each cluster;
   b) for each variable, conditional upon a proxy value equaling a mode for a cluster, assigning a value of 1 to a sub-score and multiplying an assigned value by a specified weight for the corresponding question if a weight was specified ;
   c) for each observation, summing the sub-scores for all variables to obtain an observation fitness score;
   d) summing all observation fitness scores to obtain a fitness score.

10. The method of claim 1, further comprising calculating the fitness score by activities comprising:
    a) calculating modes of proxy values for each variable for all observations associated with each cluster;
    b) for each variable, conditional upon a proxy value equaling a mode for a cluster, assigning a value of 1 to a sub-score, and conditional upon a proxy value not equaling a mode for a cluster, assigning a value of 0 to the sub-score and multiplying an assigned value by a specified weight for the corresponding question if a weight was specified;
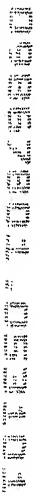
    c) for each observation, summing the sub-scores for all variables to obtain an observation fitness score;

    d) summing all observation fitness scores to obtain a fitness score.

11. The method of claim 1, further comprising modifying the data set by adding and/or subtracting observations and/or variables to improve the fitness score.

12. The method of claim 1, further comprising specifying an initial number of clusters.

13. The method of claim 1, further comprising adding a cluster to the plurality of clusters.

14. The method of claim 1, further comprising removing a cluster from the plurality of clusters.

15. The method of claim 1, further comprising obtaining a maximum number of iterations via which to arrive at a plurality of final cluster assignments that maximize fitness score.

16. The method of claim 1, further comprising identifying a maximum number of iterations to arrive at a plurality of final cluster assignments that maximize fitness score.

17. The method of claim 1, further comprising obtaining a weight to assign to a variable from the plurality of variables.

18. The method of claim 1, further comprising identifying a weight to assign to a variable from the plurality of variables.

19. The method of claim 1, further comprising re-assigning an observation to a different cluster from the plurality of clusters.

20. The method of claim 1, further comprising changing the assigned cluster for a predetermined fraction of the plurality of observations.

21. The method of claim 1, further comprising changing the assigned cluster for a predetermined fraction of the plurality of observations, said fraction including an equal number of randomly chosen observations from each of the plurality of clusters.

22. The method of claim 1, further comprising randomly changing the assigned cluster for an observation and calculating a new fitness score.

23. The method of claim 1, further comprising the activities of:

    a) changing the assigned cluster for a predetermined fraction of the plurality of observations, said fraction including an equal number of randomly chosen observations from each of the plurality of clusters;

    b) randomly changing the assigned cluster for an observation and calculating a new fitness score; and

    c) repeating activities a) and b) until cluster assignments are identified for all respondents that maximize fitness score.

24. The method of claim 1, further comprising employing linear optimization during the activities of:

a) changing the assigned cluster for a predetermined fraction of the plurality of observations, said fraction including an equal number of randomly chosen observations from each of the plurality of clusters;

b) randomly changing the assigned cluster for an observation and calculating a new fitness score; and

c) repeating activities a) and b) until cluster assignments are identified for all respondents that maximize fitness score.

25. The method of claim 1, wherein each of the plurality of observations can be initially assigned to a predetermined one of the plurality of clusters.

26. The method of claim 1, wherein each of the plurality of observations can be initially randomly assigned to one of the plurality of clusters.

27. The method of claim 1, further comprising identifying initial cluster assignments for each of the plurality of observations.

28. The method of claim 1, further comprising obtaining predetermined initial cluster assignments for each of the plurality of observations.

29. The method of claim 1, further comprising:

obtaining predetermined initial cluster assignments for each of the plurality of observations; and

modifying the data set by adding and/or subtracting observations and/or variables to improve the fitness score.

30. The method of claim 1, further comprising obtaining predetermined initial cluster assignments for each of the plurality of observations, the initial cluster assignments determined by at least one prior application of claim 1.

31. The method of claim 1, further comprising obtaining predetermined initial cluster assignments for each of the plurality of observations, the initial cluster assignments determined by a plurality of prior applications of claim 1.

32. The method of claim 1, further comprising obtaining predetermined initial cluster assignments for each of the plurality of observations, the initial cluster assignments determined by iterative applications of claim 1.

33. The method of claim 1, further comprising identifying initial cluster assignments for each of the plurality of observations, the initial cluster assignments a result of a systematic search.

34. The method of claim 1, further comprising determining initial cluster assignments for each of the plurality of observations.

35. The method of claim 1, further comprising identifying initial cluster assignments for each of the plurality of observations by the activities comprising:

identifying a pair of variables that creates a first clustering solution that maximizes the fitness score using a specified number of clusters;

determining a first single variable that creates a second clustering solution that maximizes the fitness score using the specified number of clusters;

holding the first single variable constant, determining a second single variable that, in tandem, creates a third clustering solution that maximizes the fitness score using the specified number of clusters.

36. The method of claim 1, further comprising identifying initial cluster assignments for each of the plurality of observations by activities comprising:

    a)   identifying any pair of variables that together creates a first clustering solution that maximizes the fitness score using a specified number of clusters;

    b)   determining a first single variable that creates a second clustering solution that maximizes the fitness score using the specified number of clusters;

    c)   holding the first single variable constant, determining a second single variable that, in tandem, creates a third clustering solution that maximizes the fitness score using the specified number of clusters.

    d)   holding the second single variable constant, determining a third single variable that, in tandem with the second single variable, creates a fourth clustering solution that maximizes the fitness score using the specified number of clusters;

    e)   repeating activities c) and d) by cycling through all possible combinations of variable pairings until fitness score as calculated in activity d) can be maximized.

37. A computer-readable medium containing instructions for activities comprising:

for each of a plurality of observations, obtaining a data set containing no more than one proxy value for each of a plurality of variables, each variable having a plurality of possible values;

assigning each of the plurality of observations to one of a plurality of clusters; and

for the plurality of observations and the plurality of variables, via cluster reassignment, maximizing a processor-determined fitness score representing a

number of variables from the plurality of variables for which each observation's proxy value corresponds to a mode for that observation's assigned cluster.

38. An apparatus for determining a plurality of clusters, comprising:

    for each of a plurality of observations, means for obtaining a data set containing no more than one proxy value for each of a plurality of variables, each variable having a plurality of possible values;

    means for assigning each of the plurality of observations to one of a plurality of clusters; and

    for the plurality of observations and the plurality of variables, via cluster reassignment, means for maximizing a processor-determined fitness score representing a number of variables from the plurality of variables for which each observation's proxy value corresponds to a mode for that observation's assigned cluster.